

mechanistic model [1], [3], [11], and some good reasons why the process should be unverbalizable by the subject.

11) The distinction between processes with correlates in conscious experience, and those without, merits special thought. It has been suggested [3], [7], [8] that an agent's information system needs to incorporate a "meta-organizing system" whose function is to determine goal priorities, etc., and to govern the updating operations upon the organizing system; the activity of such a meta-organizing system would seem to have some features required of the physical correlate of conscious experience [7]. On this hypothesis, the knowledge represented solely within the organizing system would be at least one category of knowledge we might "possess but not tell."

12) Fundamental to the approach advocated is the principle [1], [8] whereby concepts are internally represented in an information system by *conditional instructions for (internal or external) action*. This has recently proved fruitful in the work of Winograd [13] on the programming of computers to handle natural language, and the author believes it will reward further exploration in the present context.

13) Note that in the preceding we have been using the *converse* of the usual machine intelligence approach. Instead of starting with introspectively specified verbal performances and looking for models, we have started from a general-purpose information system equipped to develop and adapt conditional strategies to reckon with its world and asked if the mechanism of a human agent were something like this, what kinds of knowledge might he find inexpressible in words? This methodological contrast has further implications worth working out.

14) Finally, note that nothing in what has been suggested tends to support reductionist dismissals of human beings as "nothing but machines." The analysis starts from the fact that men are purposeful conscious agents. In order to retain our traditional estimate of the significance and dignity of human agency it is quite needless to foment scepticism as to its explicability, *qua* physical behavior, in mechanistic terms.

REFERENCES

- [1] D. M. MacKay, "Mindlike behavior in artefacts," *Brit. J. Philos. Sci.*, vol. II, pp. 105-121, 1951; *ibid.*, vol. III, pp. 352-353, 1953; also in *The Modeling of Mind*, K. R. Sayre and F. J. Crosson, Eds. Notre Dame, Ind.: Univ. Notre Dame Press, 1963, pp. 225-241.
- [2] —, "The epistemological problem for automata," in *Automata Studies*, C. E. Shannon and J. McCarthy, Eds. Princeton, N.J.: Princeton Univ. Press, 1956, pp. 235-251.
- [3] —, "Operational aspects of intellect," in *Mechanization of Thought Processes* (National Physical Lab. Symp. 10, 1958), Her Majesty's Stationery Office, London, 1959, pp. 37-52.
- [4] —, "Information and learning," in *Learning Automata*, H. Billing, Ed. Munich: Oldenbourg, 1961, pp. 40-49.
- [5] M. Kochen, D. M. MacKay, M. Scriven, and L. Uhr, "Computers and comprehension," Rand Corp., Memo. RM-4065-PR, Apr. 1964; also in *The Growth of Knowledge*, M. Kochen, Ed. New York: Wiley, 1967, pp. 230-243.
- [6] D. M. MacKay, "From mechanism to mind," in *Brain and Mind*, J. R. Smythies, Ed. London: Routledge and Kegan Paul, 1965, pp. 163-200.
- [7] —, "Cerebral organization and the conscious control of action," in *Brain and Conscious Experience*, John C. Eccles, Ed. New York: Springer, 1966, pp. 422-445.
- [8] —, "Ways of looking at perception," in *Models for the Perception of Speech and Visual Form*, Weiant Wathen-Dunn, Ed. Cambridge, Mass.: M.I.T. Press, 1967, pp. 23-43; also in *Perceptual Processing*, P. C. Dodwell, Ed. New York: Appleton-Century-Crofts, 1970.
- [9] —, "Recognition and action," in *Methodologies of Pattern Recognition*, S. Watanabe, Ed. New York: Academic Press, 1969, pp. 409-416.
- [10] —, "Digits and analogues," in *Principles and Practice of Bionics* (Proc. AGARD Symp. Bionics, Brussels, 1968), H. E. Gierke, W. D. Keidel, and H. L. Oestreicher, Eds. Technivision, Slough, England, 1970, pp. 457-466.
- [11] —, "The human touch," in *Pattern Recognition in Biological and Technical Systems*, O. J. Grusser, Ed. New York: Springer, 1971, pp. 20-30.
- [12] M. Polanyi, *The Tacit Dimension*. London: Routledge and Kegan Paul, 1967.
- [13] T. Winograd, "Procedures as a representation for data in a computer program for understanding natural language," Ph.D. dissertation, M.I.T., Cambridge, Mass., Feb. 1971.

Needed: A Better Theory of Cognitive Organization

GEORGE A. MILLER

Abstract—This correspondence addresses itself to the question of why computers have had so little practical success in taking over information processing tasks traditionally performed by people. This question is examined in different fields of psychology, education, and linguistics. The major reason why computers have not made more progress is that we have not had a satisfactory psychological theory of cognitive organization. Some suggestions are made as to how we might converge on a general theory of cognitive organization.

The various arguments that the human mind can accomplish things that a universal Turing machine cannot are intellectually fascinating, but as a practicing psychologist I have always been impressed by the finiteness and unreliability of the human mind; I think it might be easier to show that a universal Turing machine is far more powerful than a human mind—at least by the measures of power that are ordinarily used in these discussions. In any case, minds and brains are very different from Turing machines and computers, and the suggestion that either should properly include the other seems sufficiently absurd to justify turning to other matters.

A more interesting question is why computers have had so little practical success in taking over information processing tasks traditionally performed by people. When the computer appeared on the scene about 1950 it was obviously an answer looking for a problem, and people with problems rushed forward with claims that seemed at the time both reasonable and useful, given the obvious power that the machines promised to provide. The leaders of the new cybernetics movement issued many blank checks without much money in the bank—on principle, rather than on principal.

As we look back, therefore, it seems a bit puzzling that we have not been more successful in releasing all the intellectual power that general-purpose computers obviously could provide. One difficulty, of course, was that the power was so deliberately general purpose that a great deal of effort was required to shape the machines in the particular directions needed for particular applications. This shaping process was easier in some problem areas than others, so it is probably not surprising that the machines have been extremely successful in some fields while they have been little more than expensive toys in others. The question that I have puzzled over, therefore, is not whether the computer is limited in principle, but what has limited it in practice?

My view of this situation, of course, is conditioned by my position in the scientific spectrum. In the fields I follow most closely—psychology, education, linguistics—I believe there is some reason to feel disappointed with what has been accomplished so far. In saying this I do not wish to disparage the valiant and often brilliant efforts that have been made; when we faced the detailed problems of implementation, things simply turned out to be much more difficult than had been anticipated. My disappointment is only relative to the high hopes that were raised initially, largely on the basis of abstract considerations of what was possible in principle.

Manuscript received January 3, 1973. This work was supported by a grant from the Sloan Foundation. This paper was presented at the Workshop on Possibilities and Limitations of Artificial Intelligence, Joint National Conference on Major Systems, Anaheim, Calif., October 25-29, 1971.

The author is with the Institute for Advanced Study, Princeton, N.J. 08540.

In psychology, I think it is safe to say that the major impact of the computer has been its value as a metaphor. The realization that complex information processing could be performed by mechanical devices seemed to release many experimental psychologists from a kind of theoretical narrowness adopted as a defense against woolly mentalism. The computer metaphor freed the psychologist's theoretical imagination in a most dramatic fashion, and now the journals are full of complex information-processing characterizations of psychological phenomena. Indeed, one sometimes longs for alternatives to the steady flow of block diagrams that are currently substituting for theories. But the overall effect has been beneficial, even though it is the idea of a computer, rather than the computer itself, that seems to have been the stimulus. However, the number of actual programs that can perform complex mental tasks better than humans do is still discouragingly small, which I take as an indication that there is still much that our minds can do which we have not yet fully understood.

The computer has proved to be a valuable tool in the psychological laboratory, and a whole range of experiments is now feasible that would have been impossible or prohibitively tedious without computers.

Probably the most substantial change that computers have produced in psychology is in data processing; old methods of statistical and matrix analysis have been accelerated and made widely available by automation, and new methods of multivariate analysis have emerged that would probably have been dismissed as impractical without computers to do the work. Even here, however, progress has been much slower than one might have hoped; we have not had total systems that could integrate data collection and data analysis. Data handling, cleaning, and transforming are done in one set of programs; data analysis requires reentry with another set of programs; model building to fit the data analysis is still another set; controlling experiments based on the model requires still another. The Cambridge Project is currently investigating how all these parts could be assembled together in a computer system dedicated to psychological research, but as yet it is still a goal, not an accomplishment.

In education, of course, I need not review the evidence that indicates that first-generation applications of computers were largely unsatisfactory—I can simply refer to Oettinger's *Run, Computer, Run* for my documentation. A teacher is an intelligent educated person; to replace him with a computer is not something that we can do without very careful and extensive research on the artificial intelligence problems involved. The basic ideas of programmed instruction developed independently of computers, of course, and there is little reason to think that the art has been greatly enriched by translating the instructional programs into computer programs. There are still enormous possibilities here, just as our cybernetic fathers promised us, and I have faith that some good will come of it eventually. But the results to date have been something less than we initially hoped.

In linguistics, the major impact was felt through the financial support given to mechanical translation. It is possible that this funding served a purpose by drawing more bright young men into linguistics, but mechanical translation proved far more difficult than anticipated. I do not think that any fundamental insights into language resulted from the many hours and dollars that were spent on mechanical translation. Except for the instrumentation provided for dynamic simulations of the vocal tract, which opened up new lines in phonetic research, I am at a loss to name any advance in linguistic theory or research—and the field has been growing extremely rapidly in the last six

years—that I would attribute to the use of computers or artificial intelligence programs.

Holding this view of the computer revolution in my own field of work, I feel it is necessary to ask why things have not gone further faster.

In my opinion, the major reason we have not made more progress in these fields is that we have not had a satisfactory psychological theory of cognitive organization. That is to say, I feel sure that general-purpose computers are perfectly adequate to do anything we know how to tell them to do. The trouble is that we have not really known what to tell them to do because we do not know how we do it ourselves.

The need for a better theory of cognitive organization is probably most obvious in linguistics, where the difficulties of mechanical translation highlighted the difference between parsing and dictionary look-up, which computers can be told how to do, and interpreting and understanding messages and speaker's intentions, which we do not know how to explain to computers. However, the same obstacle can be discovered in educational applications, where the automated teacher was really stupid when measured in terms of its understanding of the subject matter it was supposed to teach. Without a better psychological theory of understanding, the computer could do little more than serve as a one-way communication channel from teacher to student. The most ambitious efforts to meet the need for a theory of cognitive organization were made in psychology, where one thinks of the General Problem Solver as an impressive pioneering effort. As in all three fields, I believe, a critically important component of any successful artificial intelligence (AI) effort has been missing. The blame, if one wishes to assign blame in such matters, lies not with the computer or with those who have tried to use it in these areas, but with the theoretical psychologist who had little more than a 19th Century conception of associationism to offer as a theory of cognitive organization.

In my opinion, therefore, before we can make any substantial progress in these branches of AI we will have to provide a better psychological theory of the mind. I doubt that such a theory can be achieved by simply locking ten psychologists in a think-tank and refusing to let them out until they provide it. The theory will emerge, if it emerges at all, from the work of those who need it most, those whose progress is most obviously blocked by the need for it.

Such a birth process is obviously inefficient, since (to mix metaphors) all these workers are sensing different parts of the elephant. Much trial and error will be required before any general theory appropriate to all our various needs can be expected to emerge. The kind of cognitive theory that a person working on question-answering machines can be expected to develop, for example, may have little relevance to the cognitive problems faced by those interested in computer-assisted instruction, or chess playing, or perceptrons, etc. Progress can be expected in all of these areas, of course, but can we really predict any convergence on the kind of general psychological theory of cognitive organization that we so desperately need?

Perhaps convergence would be accelerated if we could agree on the range of considerations that should shape such a general theory. For example, the theory should be so constructed as to accommodate the logical operations of rational thought—truth, implication, quantification, etc.—but it cannot be dominated by deductive logic because human beings are so often illogical and because the problems of inductive logic are so difficult to formulate and so important for survival. The theory must also be formulated in such a way as to admit teleological thinking, since our goals and values are so important in determining what

we think about. Probably most important as a source of evidence, the theory must accommodate linguistic processes, since so much of our time and energy is spent communicating about our conceptualizations. Finally, the theory must be formulated in such a way that it can grow, not only because children grow in their understanding of the world, but because the total implementation of any satisfactory theory is likely to be a long and arduous undertaking, and we cannot afford to delay testing it until it is completed. It takes a man at least 20 years to achieve the kind of cognitive organization we want to understand, and he is probably designed by evolution to do it rather efficiently; it would be hopelessly optimistic to think we could bring up an intelligent machine with any smaller investment of educational effort.

If one takes the developmental criteria first, then it would probably be advisable to begin with the structure of the physical world—objects, locations in space, moments in time—and give some theory of how the mind appreciates these basic coordinates of experience: how names are assigned to things, verbs to motions, and properties and relations to things in motion; how logical induction contributes to our perceptual organization of the physical world, and how deduction functions in testing the system; and how movements accomplish purposes important to the system. At the present time there are AI projects underway with many of these goals in mind; cognitive theorists should have much to learn from their successes and failures.

From this core conceptualization of physical objects located in and moving through physical space and time it might be possible to extend the theory further into more abstract and symbolic transactions. I cannot even guess how many man-years of effort would be necessary before any general theory of cognitive organization would emerge. It is possible that the outlines of such a theory are already emerging and that a major bottleneck is about to be broken. But it is also conceivable to me that this is the impossible dream of AI, that the mind can never understand itself, and that miniature theories of cognition, adequate for particular problems, are all we can ever hope to formulate.

These suggestions as to how we might converge on a general theory of cognitive organization have carried me well beyond the main point I wished to make, however. The fact is that we have not even begun to realize the intellectual potential that computers could provide, and I believe we will not make very rapid progress toward that realization until we have formulated better psychological theories of cognitive organization.

Problems in Assessing Intelligence in Artifacts

SAMUEL C. REED

Abstract—The concept of “equivalence,” as it relates to comparing man and automaton, is explored and the notions of response equivalence and process equivalence are contrasted. Some areas of psychological research are reviewed from the perspective of whether or not they concentrate on response or process identification. Finally, some questions concerning the description of biological automata are raised.

Rather than setting forth any new thoughts on the mind-body problem, or demonstrating the logical impossibility of a machine arriving at a set of results given some premises, or defining some act or process which is the sole province of man, I will briefly discuss three areas which are closely related to the central issue of whether or not a machine can be deemed intelligent: 1) “What do we mean by ‘equivalence’ between man and machine?”; 2) “What is gained through discussing ‘intelligence’ in artifices rather than in the more abundant instantiations around us?”; and 3) “To what in the field of psychology can we look in understanding the scope, complexity, and ontogeny of intelligence?”

Equivalence between man and machine is an ambiguous concept. Roughly, it means that the machine and man “do the same thing” in a particular situation. This, of course, is not sufficiently precise for many problems, and what is meant by “the same thing” may differ for different observers. The change to “evidence the same response in identical environments” does not add a great deal for reasons to be suggested later. Let us explore some of the different meanings and interpretations of the concept of “equivalence.”

The Turing criterion [1] of having a machine be behaviorally indistinguishable from a man is familiar to us all. Essentially all that is implied (not meaning to underestimate the difficulty of achieving this end) is that the machine will generate only responses which belong to the set of responses that a man would produce. This criterion leaves open the problems of defining the set of conceivable responses and establishing the criterion set against which one might judge a response. All that is required is that there be some uncertainty for the observer (presumably human) concerning the nature of that which produced the response(s).

Response equivalence in a sense analogous to the Turing criterion differs from that used in, say, learning studies. In the latter the response or responses to be considered are operationally defined *a priori*. Thus response patterns, such as bar-pressing rates in a conditioning experiment, may be relatively unambiguously compared between two groups of rats or between a group of rats and a theoretical model. Statistical tests abound for testing whether or not the performances are detectably and reliably different. While the orientation of the experimental psychologist has been principally oriented to the rejection of null hypotheses concerning means and variances, which is certainly not the same as demonstrating equivalence, some recent work in statistics and psychometrics has been directed toward the testing of hypotheses concerning the identity of measures and of performances [2], [3].

The use of a response equivalence criterion for assessing whether or not man and machine are the same is in some senses a more difficult criterion to fulfill than is the Turing criterion. First, the task of operationally defining the response variables is an enormous one. Second, the strict criterion requires that the observer have complete uncertainty after the experiment is run.

Even in the case of response equivalence one is limited with respect to the degree one may generalize about the performance of an artifice on variables not yet examined. To make any inductive leap at all, some presumption concerning the process involved in producing the responses is required. This opens the door to the area of process equivalence.

Hypothesis testing and modeling typically have as their goals the relating of input and output measures in a veridical manner. The model or hypotheses thus formed “stand for” the process being represented. Now suppose a degree of response equivalence has been established between a man and a machine in some